Determinant of Co-authorship in economics: the French case

Damien Besancenot, Kim Van Huynh, Francisco V Serranito

Economics Bulletin

Volume 35, Issue 1

Determinant of Co-authorship in economics: the French case

Damien Besancenot
Université Paris 13, Sorbonne Paris Cité – CEPN, CNRS (UMR 723)

Kim V. Huynh
Université Paris 2

Francisco Serranito
Univ. Orléans, CNRS, LEO

Abstract

This paper aims at estimating the determinants of co-authorship in economics. More specifically, we test if there is a relationship between the efficiency in research of an individual and that of his co-authors using a novel database of French academics. The main empirical result is that the number and the quality of a researcher's co-authors reflect the productivity of this researcher.

We would like to thank Kevin E. Staub for his help in programming the PQL estimation, all the participants of the 62d AFSE Meeting and the comments of the two anonymous referees. All remaining errors are ours. This research has been conducted as part of the project Labex MME-DII (ANR11-LBX-0023-01)

Citation: Damien Besancenot and Kim V. Huynh and Francisco Serranito, (2015) "Determinant of Co-authorship in economics: the French case", *Economics Bulletin*, Volume 35, Issue 1, pages 680-693

Contact: Damien Besancenot - besancenot.damien@univ-paris13.fr, Kim V. Huynh - Kim.Huynh@u-paris2.fr, Francisco Serranito - francisco.serranito@univ-orleans.fr.

Submitted: March 30, 2014. **Published:** March 22, 2015.

1. Introduction

One of the most documented trends about publication norms for the last four decades is the increase of co-authorship in the production of journals articles. A number of recent papers underline the monotonic increase in number of authors per paper. For instance, Card and DellaVigna (2013) stress thatamong the papers published in the 'top five' economic journals[1], the average number of authors was 1.3 in the early 1970s, while this number is close to 2.2 in 2011-2012. Recently, Hamermesh (2013) also notes that after the apparition in 1993 of the first four-authored papers in the American Economic Review, the Journal of Political Economy and the Quarterly Journal of Economics, the first five and six-authored papers start to appear since the beginning of the 2010's.

This increase comes with a rise in the share of co-authored papers. Ten years ago, Laband and Tollison (2000) highlighted that the percentage of co-authored papers in the previous three main economic journals increased from 10% during the 50s to 70% in 1994. More recently, Cardoso et al (2010) reported a fall in the average share of the US economists publishing singled author's papers in the top tier journals (from 45.8% during the period 1991-1996 to only 29.5% in the years 2002-2006). These recent papers confirm the results of a number of previous articles (see for instance Beaver and Rosen 1978; Stefaniak 1982; Petry 1988; Zitt et al. 2000; Beaver 2001; Glanzel 2002).

In the sociological tradition, co-authorship seems unnatural as it lessens the individual prestige of being at the origin of new analyses or newconcepts (Stephan 1996). Moreover, co-authorship requires coordinationefforts, imposes compromises between authors or may limit the innovativecontent of the papers as authors may have different degree of risk aversion(Hudson 1996).If we exclude the specific cases where collaboration is based onfriendship or is considered as a way to escape academic isolation (Medoff,2007, Acedo et al., 2006, Hamermesh, 2013), co-authorship must be explained through its positive effects on the quantity and the quality ofresearch output.Obviously, co-authorship is a simple way to increase the number of papers that a researcher may publish during a given period of time. Even if assessment procedures discount the value of coauthored papers according to the number of authors, the value of two bi-authored papers is generally worth more than the value of a single-authored paper. This creates astrong incentive to scientific collaboration (Liebowitz and Palmer, 1983, Barnett et al., 1988). Co-authoring is also an efficient way to improve the quality of a scientific production as indirectly proved by Diamond (1985) who found that the monetary return from a citation worth more to the authors when the paper is coauthored than when there is a single author.

Co-authorship has been the subject of an abundant literature, howeverlittle attention has been devoted to analyze authors' jointcharacteristics within co-authors teams. In a pioneer paper, McDowell and Melvin (1983) argue that authors, involved in increasing specialization,are looking for complementaritiesthrough co-authorship (see also Butler 2007). On the opposite, Barnett et al. (1988) and Medoff (2007) found that substitutability may be at the heart of collaborative combinations of equally skilled researchers. According to Fafchamps et al. (2006)

[1]The American Economic Review, Econometrica, the Journal of Political Economy, The Quarterly Journal of Economics and The Review of Economic Studies.

collaboration is most likely between authors of a similar level of ability (what they call assortative matching). However collaboration between authors with different abilities may also arise if the contribution of the lower ability author relaxes the time-constraint of his/her co-authors. Thus, as a result of collaboration, higher ability authors produce more research while lower ability researchers produce better quality output than would otherwise be the case.

Our paper aims at contributing to this literature by estimatingthe link between the “individual research quality” of an academic and that of his/her co-authors. Our goal is to document the fact that the higher is the skill of a researcher, the higher will be the number and the efficiency of his co-authors. The determinants of co-authorship will be estimated with a novel database considering all academic economists with aposition in a French university.

For each researcher in the dataset, and for each of their co-authors, we computed their h and g indexes. In order to measure the number and thequality of the coauthors in a one-dimensional variable, we also computed two new Meta indexes (hereafter the hh and gg indexes) built by reference to the hand the g methodologies[2]. A high hh (or gg) index reveals that an author collaborates with a high number of influential co-authors. If there is a positive link between the individual research qualities of academics within a coauthor team, then the hh index (resp. gg index) of a given author should beexplained by the author's h (resp. g) index. In a nutshell, the moretalented is a researcher, the higher will be the number and the quality ofhis/her co-authors.

The database also incorporates several control variables in order to eliminate theinfluence of side effects such as age, gender, reputation andlocalization ofthe university. Our empirical results are consistent with the assortative matching argument which leads researchers of equivalent abilities to worktogether.

The paper is organized as follows. The next section presents the empirical methodology. Section 3 describes the database; section 4 provides the results of the empirical model; and finally section 5 concludes the paper.

2- Empirical Methodology

Asa highly skilled author should have more co-authors of better quality, the goal of this section is to estimate the relationship between the “individual research quality” of an academicand that of his/her co-authors. We will then estimate a relationship as follows:

$$Q_{i,co-authors} = f\left(Q_i, W_{ij}\right), \text{ for i=1, .., N and j=1, ..., M}$$

where Q_i stands for the individual research quality of economist i, $Q_{i,co\_authors}$ represents the average quality of his co-authors and W_{ij} stands for exogenous variable j of economist i. Regressors include demographic variables, gender, academic position and a network effect. Individual performance is measured either by the h or the gindexes whereas the quality of the co-authors is defined by either the hhor the ggindexes (see the next section for further details). Both measures of the research quality of individuals are count data.

[2] Meta indexes are often used in bibliometrics. For instance, Schubert (2012) computes a Meta index to measure the characteristics of a researcher's network and Tol (2008) proposes a generalized g index to rank groups of researchers.

Poisson regression models provide a standard framework to estimate count data (seeRidout, Demétrio&Hinde, 1998, for a review). However count data suffer from two major drawbacks: overdispersion and excess of zeros. Overdispersion could stem from unobserved heterogeneity which causes the conditional variance of the sample to be larger than the conditional mean. Excess zeros arise when the number of zeros in the data exceeds what would typically be predicted by the Poisson distribution. Lambert (1992) introduced the Zero Inflate Poisson (ZIP) model in which the zeros values are the result of both a Poisson model and a Logit decision process. This model supposes that the data generating process is different for the sample values equal to zero and those which take positive value. There should be also a distinction between "structural zeros" (which are inevitable) and "sampling zeros" (which occur by chance). In order to model both unobserved heterogeneity and excess zeros a Zero Inflated Negative Binomial (ZINB) model could then be applied to the data.

The observed quality level of the co-authors ($Q_{i,co-authors}$) is modelled as function of two variables:C_i the collaboration decision variable of individual i and Q_i^*a latent variable measuring the potential quality level of the co-authors:

$$Q_{i,co-authors} = \begin{cases} Q_i^* & if \quad C_i = 1 \\ 0 & if \quad C_i = 0 \end{cases}$$

The probability function of the quality level of co-authors is then defined as follows:

$$f(Q_{i,co-authors}) = \begin{cases} p_i + (1 - p_i) \times g(0) & if \quad Q_{i,co-authors} = 0 \\ (1 - p_i) \times g(Q_{i,co-authors}) & if \quad Q_{i,co-authors} > 0 \end{cases}$$

where$p_i \in [0,1]$ is the probability that individual i will not collaborate (or the probability of a structural zero), and *g(.)* is the probability function of the parent count model. Excess zeros occur whenever $p_i > 0$. The collaboration decision C_i will depend on a new latent variable C_i^*which will be modeled with a logistic model:

$$C_i = \begin{cases} 1 & if \quad C_i^* = X_{ij}^{'} \delta_1 + \varepsilon_i \geq c_i \\ 0 & if \quad C_i^* = X_{ij}^{'} \delta_1 + \varepsilon_i < c_i \end{cases}$$

where X_{ij} are the exogenous variables involved in the decision process, c_i a threshold value and ε_i a residual following a logistic density function. Accordingly, the probability of a structural zero is then:

$$p_i = \frac{exp(X_{ij}^{'} \delta_1)}{1 + exp(X_{ij}^{'} \delta_1)}$$

A fully parametric zero-inflated model is then obtained once the probability function of the parent count model is specified. If *g(.)* is a Poisson probability function, then we get the ZIP model (Lambert, 1992):

$$\begin{cases} g(Q_{i,co-authors}, \lambda_i) = \dfrac{exp(-\lambda_i) \times \lambda_i^{Q_{i,co-authors}}}{Q_{i,co-authors}!}, \quad \lambda_i > 0 \\ \lambda_i = exp(Z_{ij}^{'} \delta_2) \end{cases}$$

where Z_{ij} are the exogenous variables explaining the expected value of the co-authors' quality level. The expected value of the quality of the co-authors is then:

$$E\left(Q_{i,co-authors}/Z_i, X_i\right) = (1 - p_i)\lambda_i = \frac{exp\left(Z_{ij}^{'}\delta_2\right)}{1 + exp\left(X_{ij}^{'}\delta_1\right)}$$

The parameters δ_2 and δ_1 can be interpreted respectively as the semi-elasticities of the parent model and the changes in the log-odds of strategic zeros. The former specification implies that each subject is observed for the same time interval, referred to as the exposure. If different subjects have different exposures (t_i), then the natural logarithm of the exposure must be included as an offset variable (Rabe-Hesketh&Skrondal, 2005):

$$\lambda_i^{'} = \lambda_i \times t_i = exp\left(Z_{ij}^{'}\delta_2 + ln(t_i)\right)$$

Parameters could be estimated either by full maximum likelihood (ML) estimation or by quasi-ML estimation. If the model is correctly specified then the ML estimators are consistent and asymptotically efficient. However, if the model is misspecified then only the quasi-ML estimates are efficient. Recently, Staub&Winkelmann (2013) propose a Poisson quasi-likelihood (PQL) estimator which is robust to misspecification of the overdispersion.

3- Description of the Database

We built an original databaseconsidering all economists with a position in a French university at the endof 2004. The information was taken from the "Tableau de classement dupersonnel enseignanttitulaire et stagiaire", economics section, NationalCouncil of the Universities (2004). This document identifies all the academics employed by French Universities on December 31th 2004 and it has the following information: gender, date of birth, academic status (Full Professor, AssistantProfessor or else)and the university assignment.

As it is now common in the literature, we will use Google Scholar citation indexes in order to compute individual research productivity. There are large empirical evidences showing that citations have a more important effect on academic earnings than the number of publication. For example, Hamermesh et al. (1982) on a sample of 148 full professors of seven large US universities prove that an increase in the total number of citations by one unit has a larger effect on academic wages than one additional published paper. So foreach academic, at the beginning of the year 2012, we downloaded his/her academic CVfrom Google Scholar (with the software PoP,"Publish or Perish", Harzing 2010) and we collected the number of citations for each paper in order to compute theauthors' h and g indexes (Hirsh 2005, Egghe 2006)[3].

[3]If these indexes provide useful tools to measure both the qualitative and the quantitative dimension of an academic research quality in a simple one-dimensional variable they are also subject to various criticisms by authors who consider that they don't correctly reflect the true influence of a given researcher (see for instance Bornmann and Daniel 2007). Moreover, it is worth noting that applyingthe PoP software may also lead to overvaluations in the number of citations: for instance when at the end of a working paper, the editor add a list of the last working papers published by his institution, then PoP considers this list as citations. To avoid such difficulties, further research will consider alternatives index of productivity.

To complete the database,we collected the year of the PhD defense and the name of the supervisor fromthe following data base “Fichier Central des Thèses” (the FrenchNational database on PhD dissertation). For foreign or not recorded thesis in the database, theinformation was obtained through individual searches on the net.

For each author we identified then the name of its co-authorsand computed their respective h and g indexes. In order to summarize boththe productivity and the number of these co-authors, wecomputed two Metaindexes denoted by hh and gg by reference to the h and the gindexes. By definition, the hh index of a researcher will be equal to n if n ofhis/her N co-authors have at least a h index equal to n, and the other (N-h) co-authors have a h index less than n. In a same way, the ggindex will be equal to n if the sum of the g indexes of his/her n bestcoauthors is superior or equal to n² (the square of the rank) and thesum of the g indexes of the n+1 best coauthors is inferior to (n+1)².These two indexes aim at giving in a one-dimensional variable a measure of both the number and the productivity of a researcher's coauthors. A high gg index mainly means that anauthor works with academics presenting a high g index (some co-authors havepublished very influential papers). A high hh index indicates that aresearcher presents a high number of productive co-authors (with high hindexes).

The raw data extracted from Google scholar presents numerous shortcomings. First it takes only an inventory of the papers contained in Google scholar and is not fully exhaustive. Moreover, authors with names identical to first names may be credited withthe work of the economists with these first names without possibility ofdisambiguation. For instance, the query "Benoit Robert" would leadindifferently to the work of Robert B or Benoit R. In the same way, authorswith frequent last name (such as Cohen) are associated to overestimatedfiles including papers from homonymous researchers. Married women who useddifferent author names during their academic life often presentedunderestimated academic resume. In such cases and for any author for whichthe disambiguation was hazardous, the author names were removed from thedatabase. From an initial total number of 1830 names in the "Tableau de classement" database we kept only 1597 researchers' names.For each author the following variables have been computed:

- A dummy variable equal to 1 if the individual has written or published at least one paper with his PhD supervisor (“COWRITE_DR”): this variable allows measuring the implicit recognition by a supervisor of the quality of his/her PhD student.
- We also control for the position of the individual in the academic career: Assistant Professor (“MCF”), Full Professor (“PR”) or other (“MA”) with dummy variables. There are three different types of position for full Professor (PR_CE, PR_1C and PR_2C) and two for Assistant Professor (MCF_HC and MCF_CN). These variables could also reveal the quality of an individual as the transition from one position to another one (say from MCF to PR or MA to MCF) is not automatic but depends on winning a contest based mostly on the number and the quality of publications[4].
- “FEMALE” is a variable equal to 1 if the individual is a woman: this variable allows taking into account a gender effect on the publishing strategy of an individual, if any.

[4]It would also be interesting to take into account the social status of the researcher during the PhD period (grant holder, scholarship holder …), but this information was not available.

- "AGE" stands for the age of the individual: publishing strategy of "young" and "old" researchers may be different and this variable allows controlling for a kind of generation effect.
- "NUMBER_YEARS" stands for the number of years since the PhD defense. This variable may be interpreted as the academic professional experience. It is worth noting that for each academic, the hh or gg indexes are computed from the beginning of his/her academic career; so the time exposure is different for each individual of the sample
- We also include interaction effects of the h (or g) index with some other exogenous variables in order to model some nonlinear effects of these exogenous variables on the quality of the co-authors. The exogenous variables retained are the following: PR and COWRITE_DR[5].
- Finally we also control for a network effect with the variable "UNIVERSITY_NAME" which is the university assignment of the individual. This is a dummy variable equal to 1 if the individual works in the assigned university. In our dataset there are 90 different institutions (universities, "Grande Ecole"). We assume that belonging to academic institutions with large economic departments which are recognized nationally and internationally can facilitate the matching with complementary co-authors[6].

In Table 1 are reported some descriptive statistics of our database. About 65% ofindividuals are Assistant Professors and 35%Full Professors. The average academic is 53 years old and he has around 22 years of professional experience. 28% of individuals in our sample are women. As regards individual research productivity indexes, 22.5% of French academics have never produced a paper referenced by Google Scholar during their career. The average French academic economist has published around 8 papers in his career. The estimated mean is equal to 2.3 for the h index and 4.8 for the g index. There is a huge heterogeneity in "quality" among French academics as the h index ranges from 0 to 24 and 0 to 68 for the g index[7]. As regards co-authorship, on average a French academic has had 4.5 different co-authors. The mean of the hh index, which summarizes the number and the quality level of the co-authors, is 3.2 and the hh index ranges from 0 to 29. There is overdispersion in the data as the hh index variance is 11.7. A similar result is obtained with the gg index with an even larger range of variation.Finally about 15% of the individuals in the sample have writing at least one paper with their PhD supervisor.

In tables 2a and 2b are reported some statistics by gender and by academic position. The most striking features are the following: on average individual research productivity indexes are lower for women than for men. Amongst Full Professors, there is an overrepresentation of men compared to women. If on average an academic has 4.5 coauthors, a Full Professor has on average 8 coauthors against 2.6 for an Assistant Professor.

[5] We have also estimate the model with other exogenous variables; but they were never significant. So we have decided for a question of saving space not to report the empirical results.
[6] As recently showed by Bosquet& Combes (2013), the network effect should be better measured at the level of the economic departments rather than universities. However in our dataset, we were unable to obtain this information.
[7]Our results are relatively in line with those obtained by Bosquet& Combes (2012) on a sample of 2,832 French economist academics for the year 2010. They obtained an average value of 2.76 for the h index and 6.96 for the g index. The discrepancy in the results may be explained by different kinds of population under study. In our sample, we only take into account academics that have a position in a French university whereas in Bosquet&Combes 's sample all full-time researchers from the CNRS (Centre National de la RechercheScientifique) and the INRA (Institut National de la RechercheAgronomique) are included. These last two kinds of academics do not have teaching loads so they may have higher research productivity.

The low value of the h and g indexes in our database reflects two opposite phenomenon. Firstly, Economic Faculties in French Universities became independent entities(from Law Faculties) at the end of 1957, but they have become autonomous in terms of recruitment of professor in economics only since the end of 1968. So, professors in economics who have been hired at this period were free of any publish or perish "diktat" and, in 2004, they were close to retirement. For these researchers, publication norms dramatically depart from the current academic standards and their h index were naturally low (the fact that most of their contributions were published in French language journalsalso contributes to a low number of citations). Secondly, our database only includes researchers appointed by a French university, it excludes a large number of prolific researchers affiliated with public research centers, such as CNRS, or employed by some "GrandesEcoles" and Business Schools institutions[8]. Finally, it is worth noting that in 2012, the software PoP offered the option to select papers according to specific subject areas. In order to limit the problems of authors' homonymy, we chose to restrict our data set to the papers identified within the option "Business, Administration, Finance, Economics" only. As a side effect, this excludes the record of many relevant works falling in the option "Social Sciences, Arts, Humanities" and contributes to a reduction of the various indexes.

4- Empirical Results

According to Staub&Winkelmann (2013), if there is misspecification of the overdispersion parameter, then the ML estimation is inconsistent and they advocate applying a PQL framework. In table 2 we have reported PQL estimates with robust standard errors for the hh and gg indexes[9].

We start by estimating the hh index determinants. Regarding the determinants of the collaboration decision, the inflate coefficient of the h variable suggests that for each unit increase in h the log odds of an inflated zero decrease significantly by 1.497 (see the inflate logit model estimates in table 2 column 1). In other words, the higher values in h the less likely the zero comes from a decision of never collaborate. This effect is nonlinear as the h squared variable is also significant. The variable "AGE" has a positive and significant (at the 1% level) effect on the decision of not collaborating as an increase in age by one year will increase the log odds of an inflate zero by 0.049. There is a gender effect in the collaboration decision as the coefficient of the female variable is positive and significant at the 10% level. Being a woman increases the log odds of being a structural zero of 0.406. Publishing with his/her PhD supervisor reduces the probability of never collaborating in the future as the coefficient estimate of this variable is negative and significant at the 10% level. More precisely, publishing at last one paper with his/her PhD supervisor reduces the log odds of being a structural zero of 0.497. The academic position does not enter significantly in the collaboration decision: the probability to collaborate is the same for a Full Professor and for an Assistant Professor.

---

[8] For instance, Jean Tirole, with a h index of 117 in 2015 according Google scholar, is excluded from the database.

[9] Identification of all parameters is achieved if at least one variable in Z is not included in X (Staub&Winkelmann, 2013). This explains why regressors in the first equation are different from those in the second one.

We turn now to comment the results for the parent model of the hh index. The "NUMBER_YEARS" variable is the offset variable as each individual has different time exposure. The h index of an individual has a positive and significant (at the 1% level) effect on the quality level of his co-authors: its estimated incidence-rate ratio (IRR) is 1.187, implying that if the h index increases by 1%, then the expected value of the hh index will increase by 18.7% (see table 2 column 1). This result confirms theoretical models stating that the quality in research of the co-authors will depend on the research quality of the author. It is worth noting that the effect of the h variable is nonlinear: the h squared coefficient is negative (-0.003=ln(0.997)) and significant at the 5%. So the impact of the h index on the hh index will be particularly strong for low values of h. Furthermore, if the individual is a Full Professor then the effect of his h index on the expected value of the hh index should be reduced by 1.1 percent point as the IRR for the H*PR variable is 0.989; but the coefficient is not significant. Thus the effect of an increase in the h variable on the expected hh-index will be the same for a Full Professor and for an Assistant Professor. So, there is no reputation effect in our sample as regards the academic position. A totally different result is obtained if the individual has published at least one paper with his PhD supervisor as the H*COWRITE_DR coefficient is negative (-0.036=ln(0.965)) and significant at the 1% level. Thus, for academicsthat have published at least one paper with their PhD supervisors, the estimated effect of the h index on their expected hh-index is reduced by 3.6 percent point. Therefore a 1% increase in the h index will increase the hh index by 15.1% for an academic who has published with his/her supervisor against 18.1 % otherwise. There is no gender effect on the publishing strategy in the data: being a woman does not reduce the expected quality of her co-authors. The estimated IRR for "COWRITE_DR" is significant and it is equal to 1.367, implying that an individual that has published at least one paper with his/her PhD supervisor has on average a 48.7% increase of the expected quality of his co-authors. However, this impact will be mitigated by the research productivity level of the young academic as the coefficient of the interaction term is negative. Therefore the most productive young researchers in French academia should avoid publishing with their PhD supervisors. As expected the AGE variable has a significant and negative impact on the quality of the co-authors: the expected change of hh index if an individual with an additional year old is -0.038 (=ln(0.963)). Compared to an Assistant Professor at the standard level (MCF_CN), being a Full Professor does not improve the hh index: indeed the coefficient albeit negative is non-significant. Oddly enough, being an Assistant Professor at the last level (MCF_HC) reduces by 13.1% the quality of the co-authors index compared to an Assistant Professor[10].

We now test for the robustness of our results by comparing the PQL estimates with the gg index as a measure of the co-authors quality. On overall the econometric results obtained with the gg index are qualitatively similar to those derived from modeling the hh index (see table 2 column 2).We get similar results regarding the collaboration decision as in the previous case. The main difference is that the gender effect is no more significant. Regarding the parent model, the G, G squared, G*COWRITE_DR, COWRITE_DR and AGE variables are significant at least at the 5% level as in the previous model.

[10]This result might stem from the fact that the transition from MCF_CN to MCF_HC should depend more on the age of an individual than on the quality of his publications in the French academia.

5- Conclusion

In this paper we test if there is a relationship between the "quality" in research of an individual and that of his co-authors using a novel database of 1597 French academics. If collaboration is a device to increase the quantity and the quality of their academic production, then efficient researchers should have more co-authors than their less productive colleagues and these co-authors should themselves be more efficient. Our empirical results confirm this intuitive argument as we estimate a positive relationship between the h index of a researcher and the hh Meta index computed with the number and the h index of his/her co-authors. As regards the others determinants of the co-authoring a paper,it appears that being a woman increases the probability of never collaborating with other economists but has no impact on the expected value of the co-authors quality. The academic position of an individual has no effect on the expected quality of co-authors. Finally, as regards young researchers, publishing with his/her PhD supervisor appears as a signal of quality in academia.

More generally, this paper states that the quality of his/her co-authors constitutes a signaling device for the quality of a researcher. In order to be fully conclusive, further research on this topic should consider additional variables to consider a wider dimension of the publication activity. For instance, besides the h and g values, the model should test alternative measures of the researchers' productivity. The size of the institution hiring the researcher, the influence of the academic network or of the research topics on the academic fellows' resume should be also considered to evaluate the robustness of our results.

References

Acedo, F.J., Barroso, C., Casanueva, C., and J. L. Galan, (2006),"Co-Authorship in Management and Organizational Studies: An Empirical andNetworkAnalysis",Journal of Management Studies, 43(5), pp. 957-983.

Barnett, A. H., Ault, R. W. and D. L. Kaserman, (1988), "The rising incidence of Co-authorship in economics: Further Evidence", Review of Economics and Statistics, 70, pp. 539-543.

Beaver, D. B. (2001), "Reflections on Scientific Collaboration (and Its Study): Past, Present and Future". Scientometrics, 52, pp. 365-377.

Beaver, D. B. and R. Rosen (1978),"Studies in Scientific Collaboration", Part I. Scientometrics 1, pp. 65-84.

Bornmann**,** L**.** and H-D Daniel, (2007)**,** **"**What do we know about the h index?**"**, Journal of the American Society for Information Science and Technology, 58(9), pp. 1381–1385.

Butler, R.J., (2007), "JRI, JF, and the internet: Coauthors, New authors and the empirical research", Journal of Risk and Insurance, 74, pp. 713-37.

Bosquet Clément andPierre-Philippe Combes, (2013), « Do large departments make academics more productive? Agglomeration and peer effects in research»,CEPR Discussion Paper n° 9401.

Bosquet Clément and Pierre-Philippe Combes (2012), « Un panorama de la recherche française en économie comparant lesapproches Google Scholar et Econlit»,*Revue d'économie politique*, Vol. 122, p. 477-545.

Card, D. and S. DellaVigna, (2013), "Nine Facts about Top Journals inEconomics," Journal of Economic Literature, 51(1), pp. 144-161.

Cardoso, A. R., Guimaraes, P. and Klaus F. Zimmermann, (2010), "Trendsin Economic Research: An International Perspective",Kyklos, 63(4),pp. 479-494.

Diamond, A. M.(1985), “The money value of citations to single-authored and multiple-authored articles”, Scientometrics, 8, pp. 315-320.

Egghe, Leo (2006), “Theory and practise of the g-index”, Scientometrics, vol.69, No 1, pp. 131-152.

Fafchamps, M., Van Der Leij M. and R. Goyal, (2006), “Scientific networks andco-authorship”, Discussion Paper Series n°256, University of Oxford.

Glanzel, W. (2002).Coauthorship Patterns and Trends in the Sciences (1980-1998): A Bibliometric Study with Implications for Database Indexing and Search Strategies. Library Trends 50(3), pp. 461-473.

Hamermesh, D., S., (2013), "Six decades of Topp Economics publishing: Whoand How?", Journal of Economic Literature, 51(1), pp. 162-172.

HamermeshD. S., Johnson G. E. &Weisbrod B. A. (1982), “Scholarship, Citations and Salaries: Economic Rewards in Economics”, Southern Economic Journal, 49, p. 472-81.

Harzing. A.-W., (2010), "The Publish or Perish Book", TarmaSoftwareResearch Pty Ltd, Melbourne, Australia, 266 p.

Hirsch, J. E. (2005). "An index to quantify an individual's scientificresearch output", PNAS 102(46), pp. 16569-16572.

Hudson, J. (1996), "Trends in Multi-authores Papers in Economics", TheJournal of Economic Perspectives, 10 (3), pp. 153-158.

Laband D., and R. Tollison, (2000), "Intellectual Collaboration", Journal ofPolitical Economy, 108, pp. 632-662.

Lambert D. (1992), “Zero-Inflated Poisson regression, with an application todefects in manufacturing”, Technometrics, vol 34, pp. 1-14.

Liebowitz, S. and J. Palmer, (1983), “Assessing Assessments of the RelativeQuality of Economics Departments”, Quarterly Review of Economics and Business28(2), 88-113.

McDowell, J.M. and M. Melvin, (1983), “The Determinants ofCo-Authorship: An Analysis of the Economics Literature”, The Review of Economics and Statistics, 65(3), pp. 155-160.

Medoff, H.M. (2007), “The Input Relationship BetweenCo-Authors in Economics. A Production Function Approach”, The American Journal of Economics and Sociology, 66(2), pp. 289-308.

Petry, G. H. (1988). "A Statistical Analysis of WorldwidCoauthorship Relationships in Scholarly Journals of Business", Journal of Economics and Business 40(2), pp. 169-176.

Rabe-Hesketh S. and A. Skrondal, (2005), Multilevel and Longitudinal Modelingusing Stata, published by Stata Press Publication, StataCorp LP, CollegeStation, Texas USA.

Ridout M., Demétrio C. and J. Hinde(1998), “Models for count data withmany zeros”, Proceeding of the XIXth International Biometric Conference, CapeTown, December, Invited Paper, pp. 179-192.

Schubert, A., (2012), "A Hirsh-type index of co-author partnership ability",Scientometrics, 91(1), pp. 303-308.

Staub E. K. and R. Winkelmann, (2013), “Consistent estimation of zero-inflated count models”, Health Economics, 22(6), pp 673-6886.

Stefaniak, B. (1982). Individual and Multiple Authorship of Papers in Chemistry and Physics.Scientometrics, pp. 4: 331-337.

Stephan, P. E., (1996), "The Economics of Science," Journal of EconomicLiterature, vol. 34(3), pp. 1199-1235, September.

Tol, R.S.J. (2008), “A rational, successive g-index applied to economicsdepartments in Ireland”, Journal of Infometrics, vol. 2, pp. 149--155.

Zitt, M., E. Bassecoulard and Y. Okubo (2000), "Shadows of the Past in International Cooperation: Collaboration Profiles of the Top Five Producers of Science", Scientometrics, pp. 47: 627-657.

Table 1: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
hh	1597	3.212899	3.418182	0	29
gg	1597	7.454602	8.58519	0	83
g	1597	4.768316	6.153543	0	68
h	1597	2.260488	2.665997	0	24
Number_papers	1597	8.364158	14.16362	0	157
Number_coauthors	1597	4.513393	7.443457	0	60
Female	1597	.280526	.4493969	0	1
Cowrite_dr	1597	.1465247	.3537423	0	1
Age	1597	52.77207	10.191	34	74
PR_CE	1597	.0563557	.2306795	0	1
PR_2C	1597	.1321227	.3387302	0	1
PR_1C	1597	.1628053	.3693035	0	1
MCF_HC	1597	.0983093	.2978257	0	1
MCF_CN	1597	.5497808	.4976715	0	1
MA	1597	.0006262	.0250235	0	1

Table 2a: Descriptive statistics by gender

gender	variable	Mean	Std. Dev.	Min	Max
Female					
	hh	2.505643	2.50553	0	12
	gg	5.744921	6.020361	0	35
	h	1.629797	1.804523	0	14
	g	3.562077	4.154004	0	26
	Number_papers	5.187359	7.561131	0	72
	Number_coauthors	3.255079	4.722665	0	37
	Full Prof.	0.1715576	0.3774217	0	1
	Ass. Prof	0.8284424	0.3774217	0	1
Male					
	hh	3.487111	3.691997	0	29
	gg	8.102222	9.350207	0	83
	h	2.514667	2.919289	0	24
	g	5.268444	6.773991	0	68
	Number_papers	9.615111	15.86342	0	157
	Number_coauthors	5.008889	8.222004	0	60

	Full Prof.	0.4168889	0.4932634	0	1
	Ass. Prof	0.5831111	0.4932634	0	1

Table 2b: Descriptive statistics by academic position

gender	**variable**	**Mean**	**Std. Dev.**	**Min**	**Max**
Ass. Prof.					
	hh	2.246334	2.342896	0	19
	gg	5.039101	5.542615	0	46
	h	1.42913	1.529428	0	16
	g	3.020528	3.75296	0	37
	Number_papers	4.473118	7.084271	0	87
	Number_coauthors	2.616813	4.186432	0	52
Full Prof					
	hh	5.018349	4.300269	0	29
	gg	11.93578	11.15934	0	83
	h	3.833028	3.541198	0	24
	g	8.100917	8.197633	0	68
	Number_papers	8.364158	14.16362	0	157
	Number_coauthors	8.073394	10.35412	0	60

Table 3: PQL Estimation Results

Dependant variable	hh		gg	
	IRR	P>lzl	IRR	P>lzl
h	1.187***	0.000		
g			1.077***	0.000
h²	0.997**	0.032		
g²			0.993***	0.000
h* COWRITE_DR	0.965***	0.000	0.986**	0.014
g * COWRITE_DR			0.986**	0.014
h*PR	0.989	0.567	0.999	0.934
g*PR			0.999	0.934
FEMALE	0.974	0.592	0.954	0.389
COWRITE_DR	1.367***	0.000	1.279***	0.003
AGE	0.963***	0.000	0.959***	0.000
Admissible Agrégation	0.980	0.890	0.971	0.836
Agrégé de Gestion	0.758***	0.009	0.764**	0.019

PR_CE	0.990	0.949	1.347	0.131
PR_1C	0.972	0.843	1.125	0.501
PR_2C	1.058	0.655	1.166	0.352
MCF_HC	0.869*	0.098	0.898	0.282
MA	1.580**	0.042	7.030***	0.000
UNIVERSITY_NAME (Network Effect)	Yes		Yes	
Inflate: Logit Model				
	Coef.	P>|z|	Coef.	P>|z|
H / G	-1.497***	0.000	-0.845***	0.005
H^2 / G^2	0.055***	0.000	0.008***	0.004
H/G *PR	0.058	0.854	0.189	0.523
AGE	0.049***	0.000	0.027**	0.032
COWRITE_DR	-0.497*	0.073	-0.811*	0.055
PR	-0.369	0.237	-0.274	0.457
FEMALE	0.406*	0.070	0.379	0.228
Constant	-1.199*	0.051	-0.342	0.616
N	1597		1597	
Log Likelihood	-2766.91		-4839.27	

The IRR value is the Incidence Rate Ratio of variable i ant it is calculated as e^{β_i}; so if regressor i is increased by 1% the dependant variable will be increased by (1-IRR) %. P-values are reported in the P>|z| column (robust standard errors have been computed). The regressor is significant at the 1% level (***), 5 % (**) and 10% (*). lnalpha indicates the overdispersion parameter of the negative binomial distribution. The offset variable is Number_Years.